



OPEN

Bioinformatics approaches identified dasatinib and bortezomib inhibit the activity of MCM7 protein as a potential treatment against human cancer

Abdus Samad^{1,2}, Md. Amdadul Huq³✉ & Md. Shahedur Rahman^{1,2}✉

Minichromosome Maintenance Complex Component 7 (MCM7) is a key component of the DNA replication licensing factor and hexamer MCM (MCM2–7) complex that regulates the DNA replication process. The MCM7 protein is associated with tumor cell proliferation that plays an important role in different human cancer progression. As the protein is highly expressed during the cancer development process, therefore, inhibition of the protein can be utilized as a treatment option for different human cancer. However, the study aimed to identify potential small molecular drug candidates against the MCM7 protein that can utilize treatment options for human cancer. Initially, the compounds identified from protein-drugs network analysis have been retrieved from NetworkAnalyst v3.0 server and screened through molecular docking, MM-GBSA, DFT, pharmacokinetics, toxicity, and molecular dynamics (MD) simulation approach. Two compounds namely Dasatinib (CID_3062316) and Bortezomib (CID_387447) have been identified throughout the screening process, which have the highest negative binding affinity (Kcal/mol) and binding free energy (Kcal/mol). The pharmacokinetics and toxicity analysis identified drug-like properties and no toxicity properties of the compounds, where 500 ns MD simulation confirmed structural stability of the two compounds to the targeted proteins. Therefore, we can conclude that the compounds dasatinib and bortezomib can inhibit the activity of the MCM7 and can be developed as a treatment option against human cancer.

Abbreviations

AA	Amino acid
CAAD	Computer-aided drug design
DFT	Density Functional Theory
GLOBOCAN	Global Cancer Observatory
η	Hardness
LYP	Lee, Yang, and Parrs
RO5	Lipinski rule of five
MCM7	Minichromosome Maintenance Complex Component 7
MDS	Molecular dynamic simulation
MM-GBSA	Molecular Mechanics-Generalized Born Surface Area
PDB	Protein Data Bank
P-L contact	Protein–Ligand contacts
QM	Quantum mechanical
Rg	Radius of gyration
RMSD	Root-mean-square deviation

¹Department of Genetic Engineering and Biotechnology, Jashore University of Science and Technology, Jashore 7408, Bangladesh. ²Bioinformatics and Microbial Biotechnology Laboratory, Department of Genetic Engineering and Biotechnology, Jashore University of Science and Technology, Jashore 7408, Bangladesh. ³Department of Food and Nutrition, College of Biotechnology and Natural Resource, Chung-Ang University, Anseong, Gyeonggi-do 17546, Republic of Korea. ✉email: amdadbge@gmail.com; ms.rahman@just.edu.bd

RMSF	Root-mean-square fluctuation
SID	Simulation Interaction Diagram
S	Softness
SASA	Solvent-accessible surface area
TSA	Trichostatin A

Nowadays cancer has been remarked as the most life-threatening disease in the world. Global Cancer Observatory (GLOBOCAN) a tool that predicts future cancer incidence, reported that 19.3 million new cancer cases in 2020 and almost 10.0 million cancer-related deaths worldwide¹. Among them, the most deadly cancers were lung cancer that affects 1.8 million people around the world and causes deaths approximately 18%. Where colorectal, liver, stomach, and female breast cancer took around 9.4%, 8.3%, 7.7%, and 6.9% live worldwide¹. The mortality rate is still increasing around the world due to aging and populations^{2,3}. Despite extensive research and improvement of therapeutic approaches, the treatment of cancer is often diagnosed at a late stage. Early-stage detection of human cancer can improve the therapeutic option resulting decrease in the mortality rate^{4,5}. Therefore, it is urgent to develop an effective therapy or drug candidates against different human cancer.

MCM7 is a part of the DNA replication licensing factor and hexamer MCM (MCM2–7) complex that regulates DNA replication^{6–9}. The hexameric MCM protein forms a double trimeric complex along with the MCM4 and MCM6 that help to the unwinding of the DNA strand resulting initiation of DNA replication^{8,10}. The frequency of chromosome breaks going to higher in cells under the replication stress due to the suppression of the MCM complex⁸. Downregulation of any one of its subunits destabilized the MCM complex and cells undergo limited replication. During the process, cells become hypersensitive to DNA replication stresses resulting in DNA damage and further inhibition of cell growth via activating the checkpoint signals^{8,11}. Recently researcher has been found that MCM7 regulates the binding activity of MCM proteins that are highly associated with tumorigenesis and promotes cancer progression^{12–15}. The MCM7 mRNA expression is a prognostic biomarker in ovarian, lung, and colorectal cancer and a tremendous biomarker in cervical cancer¹². As the MCM7 overexpression plays a crucial role in cancer development, the study aimed to identify potential drug candidates against the protein to treat human cancer.

The conventional process of developing new drugs usually involves lengthy, expensive, and requires intense effort^{16,17}. On the contrary, the computational design of a drug is relatively easier, requires low time and less effort¹⁸. The in silico virtual screening process help to generate lead compounds in the way of faster time and a lower cost¹⁹. Moreover, the computer-aided drug design (CADD) by using virtual screening process includes docking, absorption, distribution, metabolism, and excretion (ADME), toxicity, and molecular dynamics (MD) simulation has been applied to identify a diverse range of promising drug candidates²⁰.

Recently researchers have been identified trichostatin A (TSA) as a promising drug candidate against human glioblastoma²¹, gastric, ovarian, breast, cervical, small-cell lung, and gastric cancer cell²². TSA is not just a possible inducer of apoptosis yet in addition engaged in the regenerated ability of stem cells²³. Therefore, the study aimed to computationally discover potential drug candidates targeting the protein.

Initially, the study assess the protein–drugs interaction. Then targeting the protein potential drugs candidates has been identified by using the virtual screening, molecular docking, molecular mechanics-generalized born surface area (MM-GBSA), density functional theory (DFT), ADMET, and dynamic simulation approaches.

Results

MCM7 protein interaction with drugs or chemicals. The protein–drug interaction of all the MCM7 proteins was mapped out through the NetworkAnalyst website. A total of 41 compounds were found as a potential inhibitor against MCM7 in human cancer (Fig. 1) and their compound identification number (CID) has been retrieved from the PubChem website in SDF file format (SM 1).

MCM7 protein–chemical compounds binding scores and MM-GBSA studies. Molecular docking between targeted protein and retrieved 41 ligand molecules has been performed by the glide v-8.8 tool. We find a total of 18 compounds having the highest binding affinity against MCM7 protein compared to the control ligand (TSA). Among 18 ligands CID_123917 shows high negative binding affinity and all docking scores range between -6.265 and -4.536 kcal/mol and the control (TSA) ligand binding score was -4.529 kcal/mol (Fig. 2 and SM 2). High 18 binding scores containing ligands were selected for MM-GBSA analysis compared to control ligand (TSA). In the MM-GBSA calculation, CID_9874191 and CID_5757 produced the negative highest and lowest MMGBSA ΔG Binding (NS) score of -71.85 kcal/mol and -21.25 kcal/mol. Among 5 ligands was showed a good binding free energy score (-71.85 to -50.65 kcal/mol) compared to control (-50.11 kcal/mol) (Fig. 2 and SM 2). Top 5 and control compounds based on their docking and MM-GBSA score were chosen for further evaluation.

MM-GBSA studies of selected compounds. MM-GBSA calculation of the selected compounds CID_9874191, CID_208908, CID_387447, CID_5330286, CID_3062316, and CID_444732 (control) produced the negative MMGBSA ΔG Bind(NS) score of -71.85 , -60.32 , -56.97 , -54.41 to 71.85 , -50.65 , and -50.11 kcal/mol, respectively. Furthermore, examination of binding free energy values for each MCM7–ligands docked complexes significantly exposed the performance of ΔG Bind Coulomb (Coulomb energy), ΔG Bind H-bond (Hydrogen bond energy), ΔG Bind Lipo (Lipophilicity energy), and ΔG Bind vdW (Van der Waals interaction energy) in the respective complex stability. These outcomes measured the strong binding affinity of target compounds such as CID_9874191, CID_208908, CID_387447, CID_5330286, CID_3062316, by comparison to

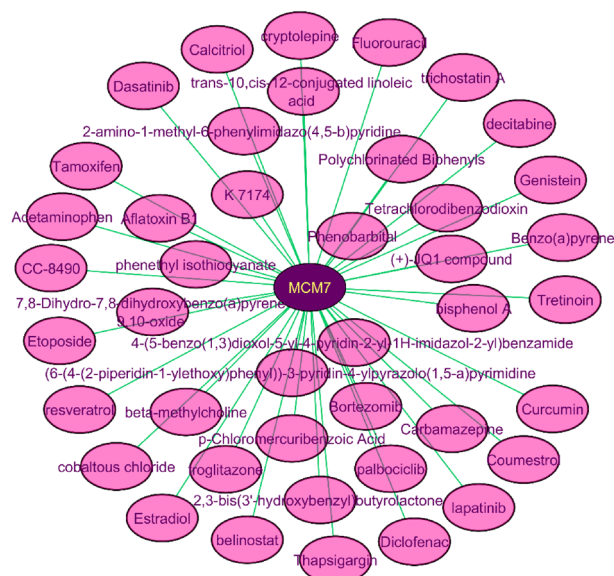


Figure 1. The protein–drugs interaction network of MCM7, where purple color represents the MCM7 protein and pink color represents the drugs.

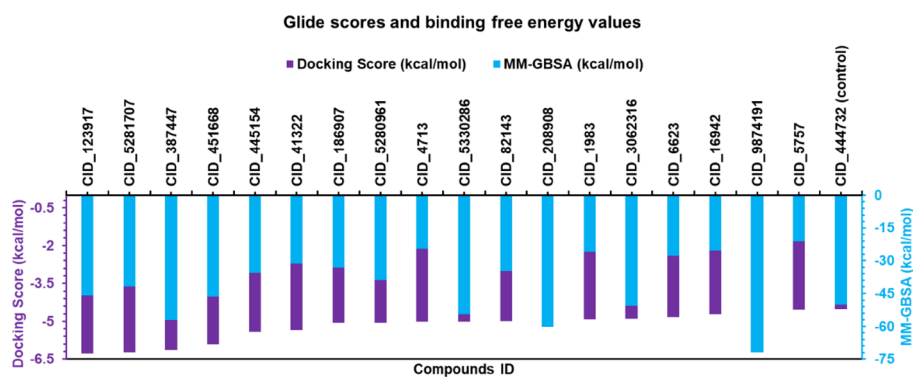


Figure 2. Bar graph representing the docking score and binding free energy values of the 18 hit compounds along with the control compound TSA.

CID_444732 (control) of human MCM7 protein (Fig. 3). The top five compounds include control compounds have been chosen based on their docking and MM-GBSA score and rederived for further analysis.

Ligand optimization by QM (quantum mechanical) calculation. In the DFT calculations, CID_9874191, CID_208908, CID_387447, CID_5330286, CID_3062316, and CID_444732 (control) generated a HOMO and LUMO energy score of -0.26642 and -0.1060 , -0.19453 and -0.06287 , -0.23891 and -0.07207 , -0.28103 and -0.14699 , -0.21012 and -0.05364 , and -0.19369 and -0.04438 a.u (Fig. 4 and Table 1).

CID_387447 and CID_5330286 generated HLG, hardness, and softness energy of 4.20262 and 1.67697, 2.10131 and 0.83847, and 0.4758 and 1.1926 eV, respectively (Table 1). Also, CID_208908 and CID_3062316 produced the HLG, hardness, and softness energy of 3.09945 and 4.07473, 1.549725, and 2.037365, and 0.6452 and 0.4908 eV, respectively (Table 1). On the other hand, CID_9874191 and CID_444732 (control) gave HLG, hardness, and softness energy of 4.04001 and 3.94145, 2.020005 and 1.970725, and 0.4950 and 0.5074 eV, respectively (Table 1). CID_387447, CID_3062316, and CID_9874191 compounds were selected compared with CID_444732 (control) compound for further analysis.

Molecular features of the selected chemical compounds. ADMET properties are one of the valuable parameters for designing effective drug candidates against a specific target. Drug likeness properties of selected compounds have been accessed according to the Lipinski rule of five (RO5). The selected five compounds have the highest negative binding affinities and obtained the RO5. ADMET properties of selected compounds have been presented in Table 2. According to the SwissADME and PKCSM server, all the selected ligands have low

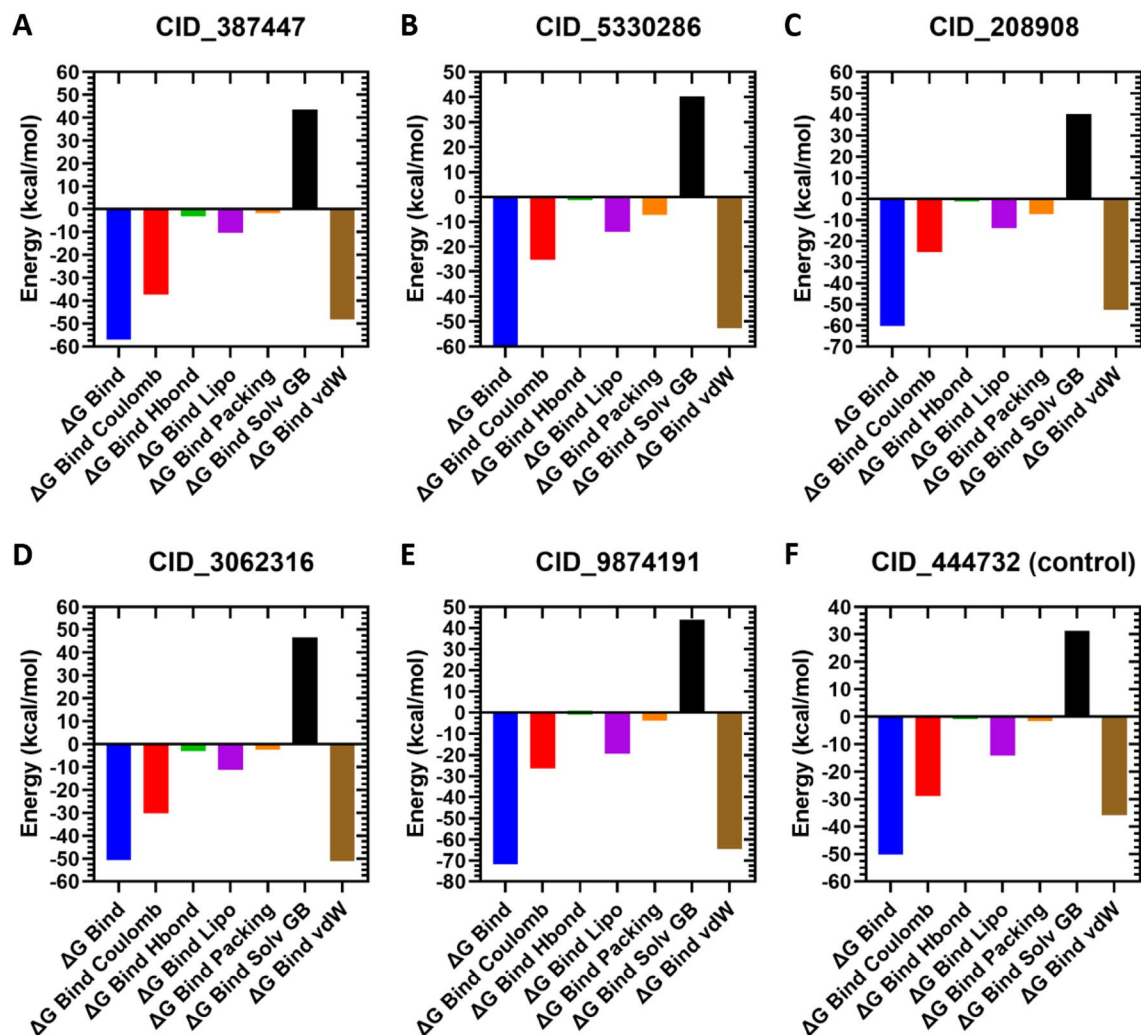


Figure 3. MM-GBSA binding free energy values of the selected docked compounds were comparison with TSA.

toxicity, no Ames toxicity, good GI absorption, good distribution rate, good water solubility (logS), and logP value expect for the compounds CID_9874191 which has a high molecular weight (Table 2).

Visualization of selected chemical compounds-MCM7 protein. The best docking, MM-GBSA, DFT, and ADMET score containing selected five compounds (CID_387447, CID_3062316, CID_9874191, and CID_444732 (control)) has been retrieved for further analyzed and molecular interactions have been visualized through Maestro v12.5 (Fig. 5, Table 3, and SM3). Different types of non-bonded interactions between receptors and ligands like hydrogen bonds, electrostatic bonds, and hydrophobic bonds have been identified. Among the selected drugs, CID_387447 showed the -6.137 kcal/mol binding affinities with MCM7 and interacted H-bond binding residues at Gly268, Leu436, Ser401, and Asn307 residue position (Fig. 5 and Table 3). The molecule of the binding affinities (-4.905 kcal/mol) has CID_3062316 formed three hydrogen bonds with the receptor and interacted binding residues was Asp395, Gly268, and Lys305 (Fig. 5 and Table 3). The CID_9874191-receptor complex (binding affinity: -4.578 kcal/mol) was stabilized with two hydrogen bonds (Lys305 and Ser401). CID_444732 (control)-receptor complex having a binding energy of -4.529 kcal/mol was stabilized with one hydrogen bond (Lys308). The binding interactions of top molecules have presented in Table 3 and illustrated in Fig. 5.

Stability of MCM7 protein-chemical compounds. The root means square deviations (RMSDs) of Ca atoms have been computed for the compounds CID_387447, CID_3062316, CID_9874191, and CID_444732 (control) protein-ligand complexes to measure the protein structure stability throughout the 500 ns simulation time. The compound having a CID_387447 complex becomes stable after 370 ns then it shows very low fluctuation 0.365 Å was found during 307–500 ns simulation time. The average, highest, and lowest RMSD value was found with a value of 2.641 Å, 3.196 Å, and 0.685 Å during 500 ns simulation time. Where the ligand average RMSD value was 2.404 Å during the 500 ns simulation time, which is the 2nd best compound compared to the

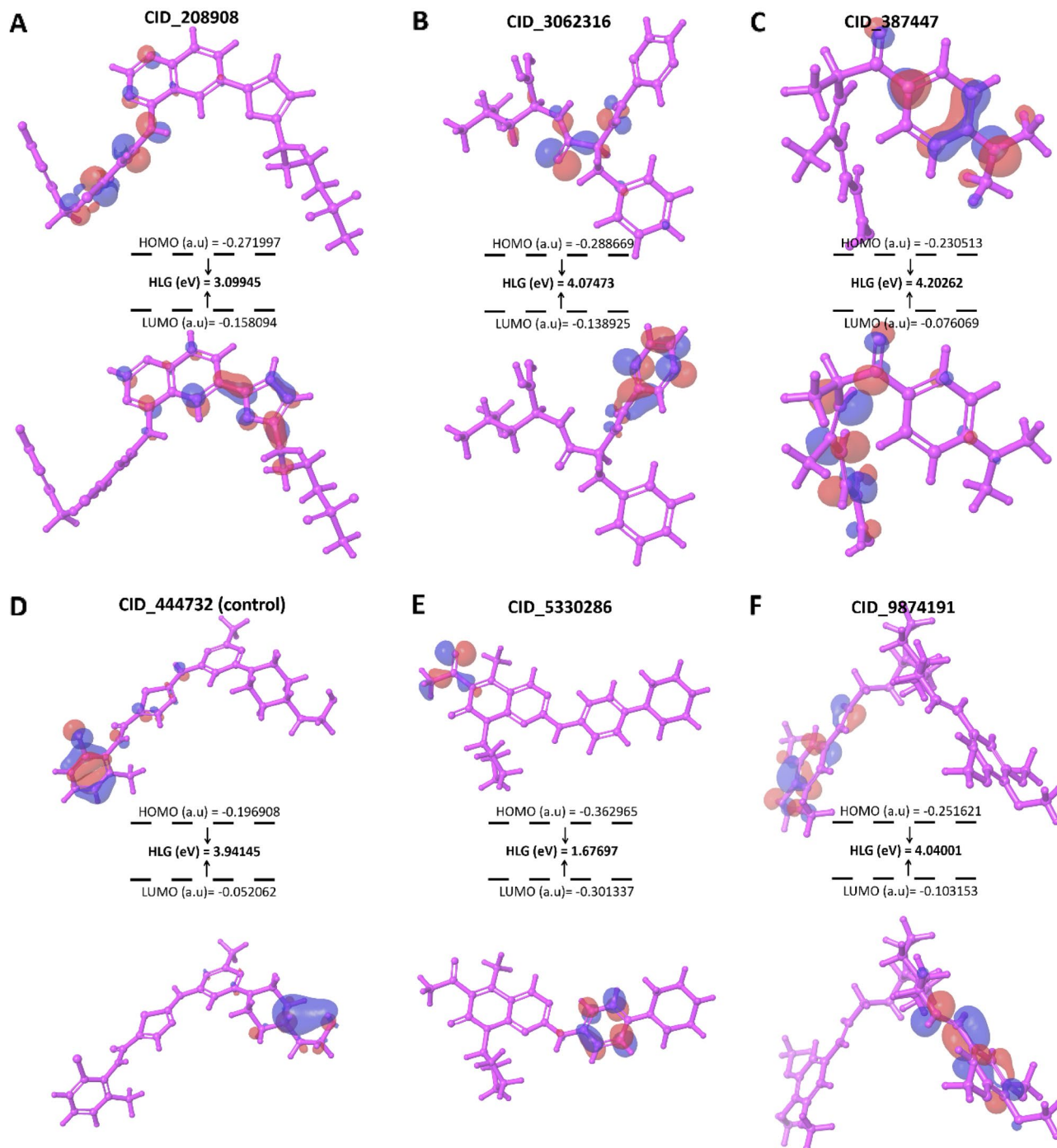


Figure 4. The HOMO and LUMO energy score and structure of the 6 docked compounds were compared with TSA.

Pubchem ID	HOMO (a.u)	LUMO (a.u)	HLG (eV)	Hardness (eV)	softness (eV)
CID_387447	- 0.23891	- 0.07207	4.20262	2.10131	0.4758
CID_5330286	- 0.28103	- 0.14699	1.67697	0.83847	1.1926
CID_208908	- 0.19453	- 0.06287	3.09945	1.549725	0.6452
CID_3062316	- 0.21012	- 0.05364	4.07473	2.037365	0.4908
CID_9874191	- 0.26642	- 0.10601	4.04001	2.020005	0.4950
CID_444732 (control)	- 0.19369	- 0.04438	3.94145	1.970725	0.5074

Table 1. DFT calculation of the selected compounds after molecular and MM-GBSA analysis.

Drug properties	CID_387447	CID_3062316	CID_9874191	CID_444732 (control)
Physicochemical properties				
MW (g/mol)	384.24	488.01	568.74	302.37
Heavy atoms	28	33	41	22
Arom. heavy atoms	12	17	12	6
Rotatable bonds	11	8	16	7
H-bond acceptors	6	6	8	3
H-bond donors	4	3	0	2
Lipophilicity				
Log Po/w	0.22	2.8	5.15	2.37
Water solubility				
Log S (ESOL)	- 2.71	- 4.98	- 6.41	- 3.19
Pharmacokinetics				
GI absorption	High	High	High	High
BBB permeant	No	No	No	No
Drug likeness				
Lipinski, Violation	Yes,0	Yes,0	Yes,1	Yes,0
Toxicity				
AMES toxicity	No	No	No	No
Oral Rat Acute Toxicity (LD50) (mol/kg)	1.868	2.676	2.986	1.897
Skin Sensitization	No	No	No	No

Table 2. Molecular features of the selected chemical compounds.

rest of the three compounds (Fig. 6A). However, the lowest RMSD was 1.2 Å during the 150 ns simulation time. The compounds CID_3062316 complex show stability after 307 ns simulation time, where its average fluctuation of RMSD 0.331 Å was found during 307–500 ns simulation time and the average protein and ligand RMSD was 2.638 Å and 2.049 Å. Herein, the compound with CID_3062316 also shows a good RMSD and the best lowest difference RMSD between protein and ligand complex system (Fig. 6B). During the whole simulation time, the CID_9874191 shows the highest and average RMSD (3.950 Å and 2.699 Å) and it is quite unstable and the average fluctuation was 0.833 Å during the 410 ns to 500 ns. The compound CID_9874191 also shows an unacceptable RMSD value and very high fluctuation (Fig. 6C). However, 444,372 (Control) complexes become stable at 100–132 ns but it fluctuates (0.492 Å) found higher after 110 ns time. At the end of the simulation, it shows 2.941 Å average RMSD which value is high compared to CID_3062316 and CID_387447 (Fig. 6D). However, the compounds CID_3062316 and CID_387447 show less fluctuation among all the four selected compounds shown in Fig. 6.

Root mean square fluctuation (RMSF) indicate the fluctuation of amino acid (AA) residue in protein structure. The high RMSF value of the amino acid residues determines the fluctuation and stability level of AA residues in a complex system. Initially, the compound CID_9874191 complex shows the highest amino acid fluctuation 3.703 Å, where all the complexes are bound to < 3 Å RMSF value. Secondly, the compound CID_444372 (Control) shows that the highest RMSF value of 4.088 Å. CID_387447 and CID_3062316 complex was showing the highest amino acid fluctuation of 2.682 Å and 2.489 Å. During the RMSF analysis, the compounds CID_3062316 and CID_387447 show the highest stability compared to control compounds depicted in Fig. 7.

To determine the protein mobility and rigidity, the radius of gyration (Rg) of the protein–ligand complex has been analyzed. The compounds CID_9874191 and CID_444372 (control) exhibited the maximum and minimum trend of fluctuations range 9.128 Å to 5.776 Å (difference 3.352 Å) and 5.199 Å to 4.078 Å (difference 1.121 Å). The CID_3062316 and CID_387447 showed the minimum and maximum fluctuations rate of 4.459 Å to 6.372 Å (difference 1.913 Å) and 3.372 Å to 4.521 Å (difference 1.149 Å), and both compounds were more stable compared control ligand (Fig. 8A).

Additionally, the Solvent Accessible Surface Area (SASA), Polar Surface Area (PSA), Hydrogen Bond (H-Bond) of the compounds have been analyzed for all the proteins and protein–drug complexes. SASA analysis is useful in understanding the solvent-like behavior (hydrophilic or hydrophobic) of a protein molecule as well as protein–ligand complexes. The compounds CID_387447, CID_3062316, CID_9874191, and CID_444372 exhibited the minimum and maximum trend of fluctuations range 507.866–53.001 Å², 140.411–459.061 Å², 135.384–560.323 Å², and 11.671–150.234 Å² (Fig. 8B). The PSA and hydrogen bond analysis of CID_3062316 and CID_387447 found the compounds have good stability (Fig. 8C,D). The results suggested that both CID_3062316 and CID_387447 protein–drug complexes were impressively stable after the binding of drug molecules.

Protein–ligand contact analysis. The complex structure of the protein with the selected ligands and, their intermolecular interactions has been evaluated at 500 ns simulation time via the simulation interactions diagram (SID). Depending on some parameters including hydrogen bond, non-covalent bond (hydrophobic bond), ionic bond, and water bridges bond the contact between protein and ligands complex structure includes

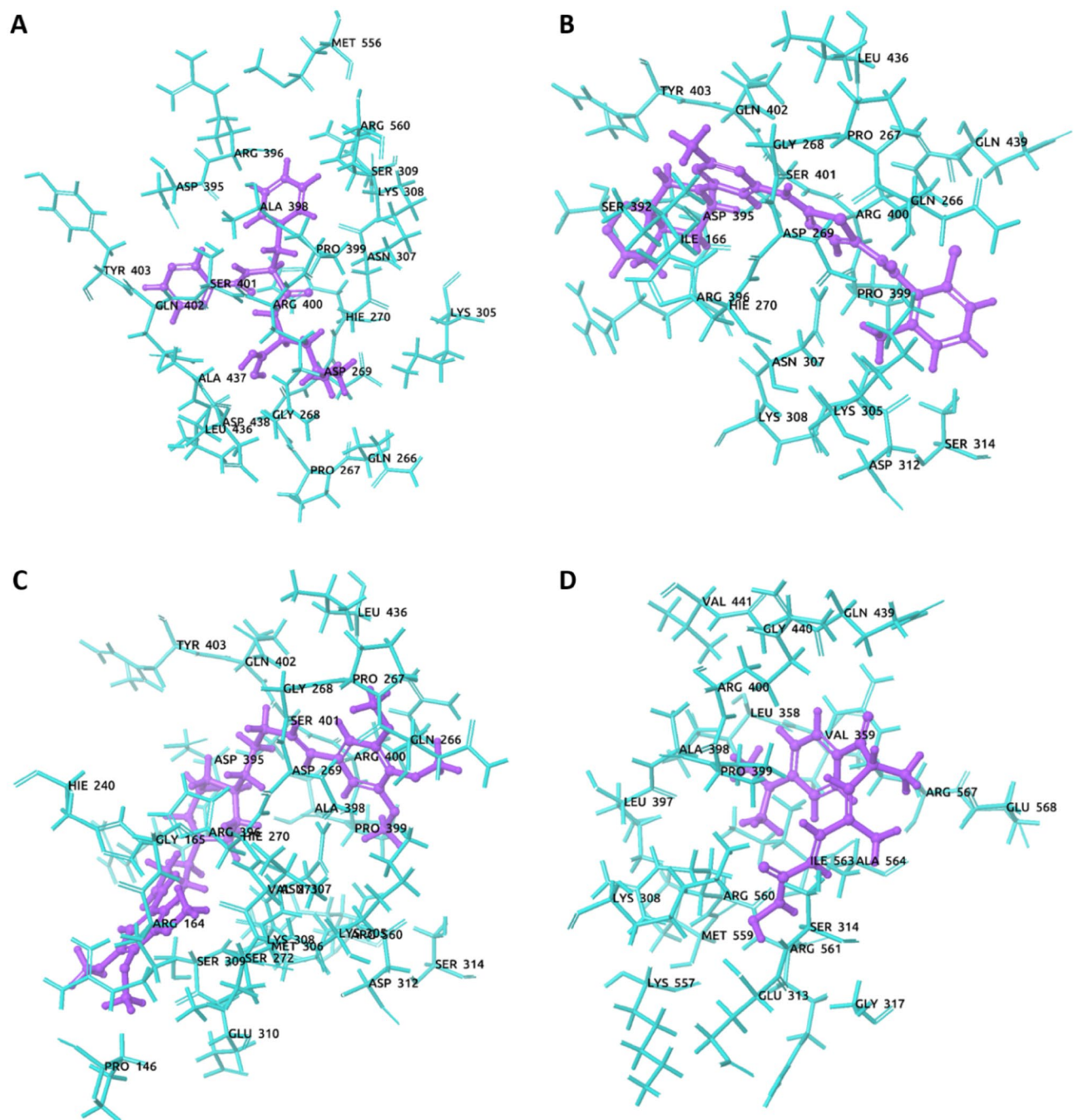


Figure 5. The 3D interaction between MCM7 protein and selected compounds complex structure. (A) CID_387447, (B) CID_3062316, (C) CID_9874191, and (D) CID_444732 (control).

Compounds	H-Bond	Other interaction
CID_387447 (Bortezomib)	Gly268, Leu436, Ser401, and Asn307	Tyr403, Gln402, Arg400, Pro399, Ala398, Arg396, Asp395, Lys308, Asn307, Lys305, Gln266, Asp269, Hie270, Ala437, and Arg560
CID_3062316 (Dasatinib)	Asp395, Gly268, and Lys305	Ile166, Hie270, Asp269, Gln266, Asn307, Lys308, Ser314, Asp312, Tyr403, Gln402, Ser401, Arg400, Pro399, and Arg396
CID_9874191 (K-7174)	Lys305 and Ser401	Met306, Asn307, Lys308, Ser309, Arg560, Arg164, Gly165, Hie240, Asp395, Arg396, Ala398, Pro399, Arg400, Gln402, Ser272, Val271, Hie270, Asp269, Gly268, Pro267, Gln266, and Leu436
CID_444732 (control)	Lys308	Gly440, Gln439, Ser314, Glu313, Leu358, Leu397, Ala398, Pro399, Glu368, Arg567, Ala564, Ile563, Arg561, Arg560, Ser314, and Glu313

Table 3. A list of various interactions and interacting residues of MCM7 with the compounds were logged from respective docked complexes.

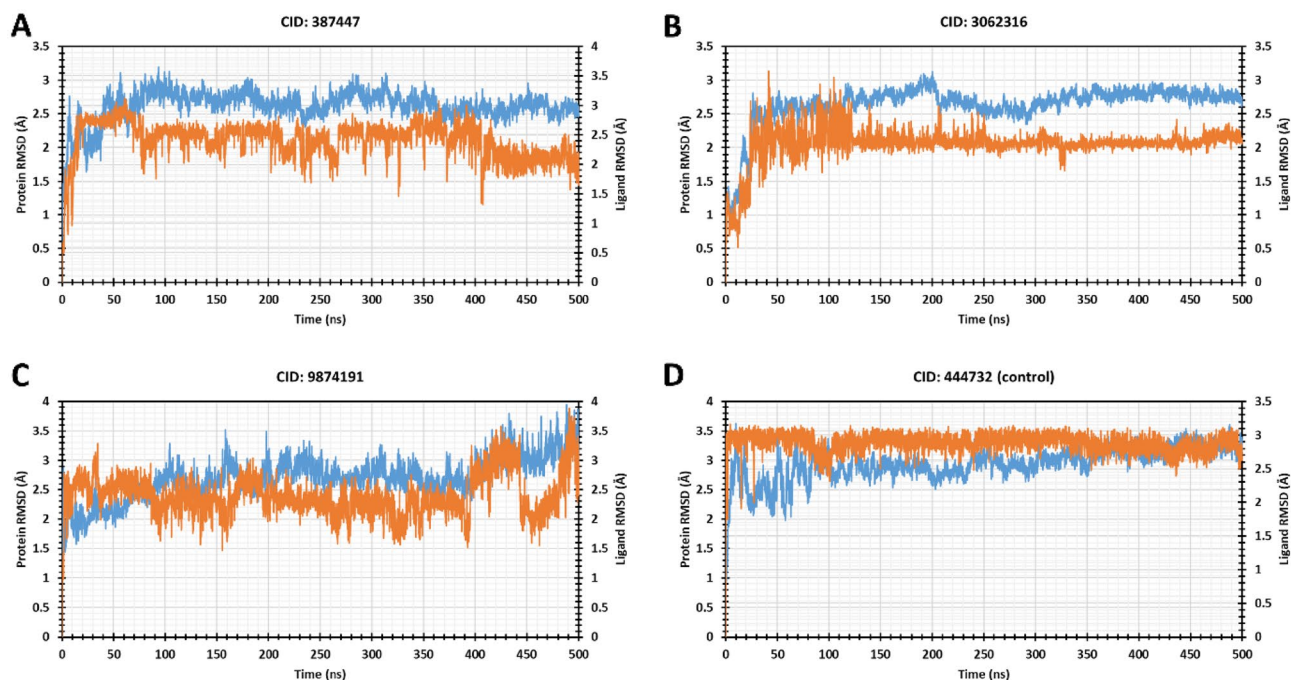


Figure 6. Graphs represent the MD simulation for the selected protein-ligands complex during 500 ns simulation time. Herein, representing the compounds (A) CID_387447, (B) CID_3062316, (C) CID_9874191, and (D) CID_444372 RMSD values of MCM7 protein (blue curves) and ligands (red curves).

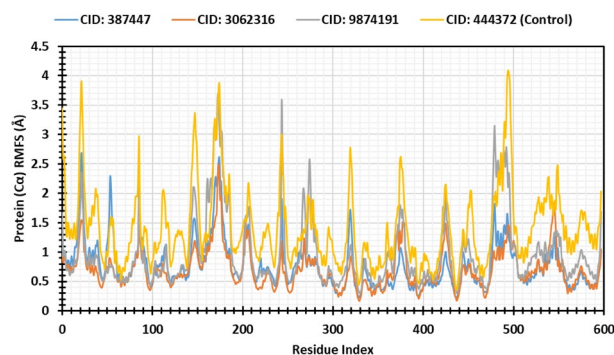


Figure 7. Graphs exhibiting the information about the RMSF values of MCM7 protein by 500 ns time of MD simulation.

CID_387447, CID_3062316, and CID_9874191 have been analyzed and represented in Fig. 9. The compound CID_387447 generated multiple (more than two) interactions at THR168, SER241, ARG396, PRO399, SER401, GLN402, and LEU436 residues with an interaction fraction (IF) value 0.3, 0.6, 0.8, 0.18, 0.5, 0.9, and 0.6 simulation time the specific interaction is maintained by the multiple contacts of the same subtype with the ligand accordingly (Fig. 9A). The compound CID_3062316 formed multiple interaction at LYS305 (0.8), ASN307 (1.2), PRO399 (0.9), SER401 (1.0), GLN402 (0.8), and LEU436 (0.4) residues maintained by simulation time accordingly (Fig. 9B). CID_9874191 compounds formed multiple interaction at ARG150 (0.5), THR168 (0.5), HIS240 (0.8), ASP395 (1.0), SER401 (0.1), and GLN402 (0.3), residues maintained by simulation time accordingly (Fig. 9C). In the case of the compound CID_444372, it has found to form multiple interactions at the position of GLU318 (0.8), PRO399 (0.6), and ARG561 (2.0) suggests that of the simulation time the specific interaction is maintained and helped to make a stable binding with the desired protein (Fig. 9D). However, the compounds CID_3062316 and CID_387447 show good hydrogen and other bond interaction with all the four selected compounds shown in Fig. 9.

Ligand–protein contact analysis. The selected three ligands CID_387447, CID_3062316, CID_9874191, and CID_444372 (control) with protein interactions have been monitored throughout the SID. The compound CID_387447, CID_3062316, CID_9874191, and CID_444372 (control) generated multiple (more than two) interactions residues of the simulation time the specific interaction is maintained by the multiple contacts of the

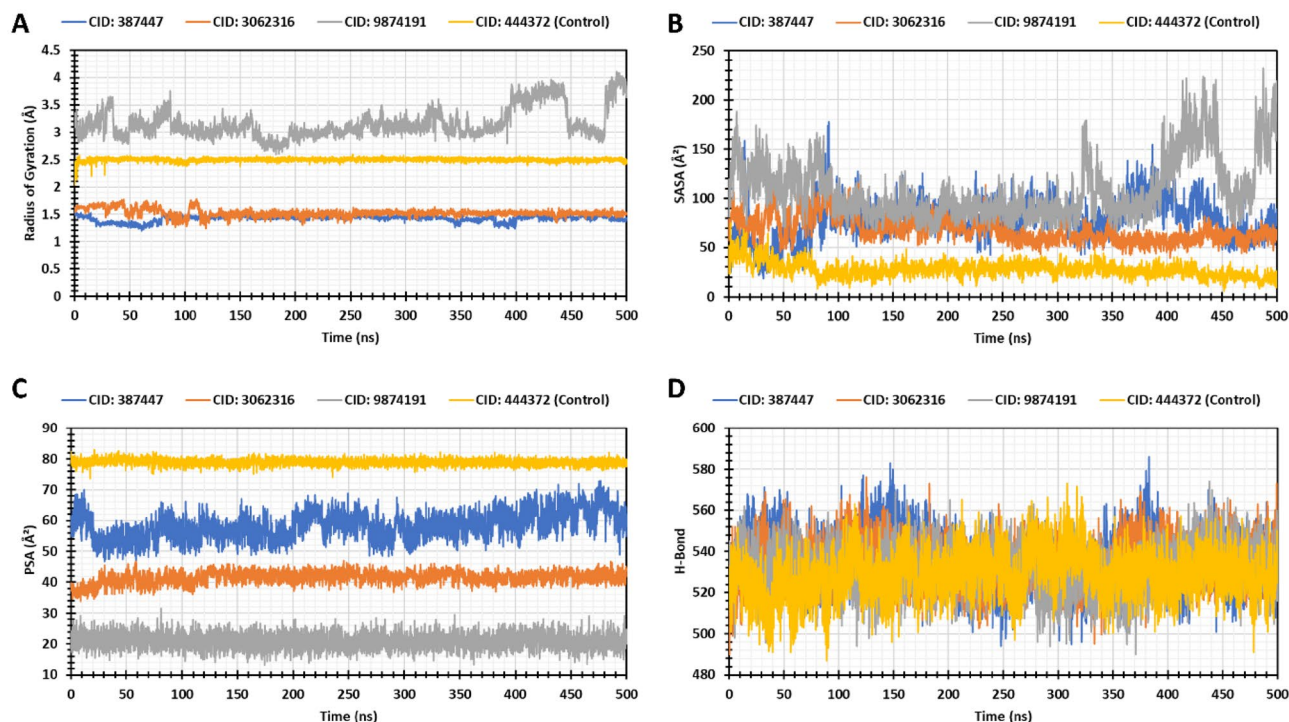


Figure 8. Graphs represent the MD simulation result by 500 ns time. Here, (A) exhibited the result of the radius of gyration (Rg) values of MCM7 protein, (B) denote the values of SASA, (C) PSA, and (D) Hydrogen bonds.

same subtype with the ligand accordingly (Fig. 10). In ligand–protein interaction analysis, CID_3062316 and CID_387447 were both compounds that are more stable compared to control compounds (Fig. 10).

MM-GBSA analysis from post molecular dynamic simulation trajectory. MM/GBSA methods have been used in this study to estimate the ligand-binding free energy to the desired protein. The MM/GBSA of the protein–ligand complex structure has been calculated from the few snapshots (~ 500 ns) of the dynamic simulation trajectory. The analysis of the complex structure found higher net negative binding free energy values -52.851 kcal/mol, -48.241 kcal/mol, -67.541 kcal/mol, and -47.107 kcal/mol for the selected four compounds CID: 387,447, CID: 3,062,316, CID: 9,874,191, and CID: 444,732, respectively with the targeted protein (Fig. 11). Therefore, it can be considered that the selected compounds will be able to maintain a long-term interaction with the desired MCM7 protein.

Discussion

Cancer is the subsequent driving reason for the highest mortality around the world. Generally, the pervasiveness of disease has expanded compared to the previous era^{24,25}. In this manner, malignant growth is a major issue influencing the wellbeing of humans. Shockingly, it is an assortment infection at the tissue level and this assortment is a significant test for its particular treatment, trailed by the ability of therapy^{26,27}. The MCM7 association in human cancer significantly immunostaining for human cancers: colorectal adenocarcinoma, endometrial carcinoma, esophageal adenocarcinoma, melanoma cancer, oral squamous cell carcinoma, thyroid cancer, and glioblastoma²⁸. In this study, we demonstrated that MCM7 mRNA expression was markedly down or up-regulated in human cancer. To date our knowledge, this is the comprehensive computational analysis conducted for the identification of potential drugs like candidates against human cancer by targeting the MCM7 protein.

Computer-Aided Drug Design (CADD) is one of the most promising tools for the selection of novel compounds against a specific protein as it includes different advanced features and techniques²⁹. The CADD approach has minimized the required time and costs involved in the entire drug discovery process that makes the virtual screening process includes molecular docking, MM-GBSA, DFT, MD simulation, and ADMET, etc. as integral parts of drug designing³⁰.

In this study, we identified MCM7 and their potential drug candidates by molecular docking and other processes. Initially, the molecular docking process has used to screen the compounds, where the top 18 compounds have been selected with the highest negative binding affinities compared to the control ligand (TSA). All the selected ligands have higher negative binding affinities than previously reported TSA compounds related to human cancer^{21,22}. In the MM-GBSA study, the most negative ΔG_{Bind} score (the lowest score) is considered as the best ΔG_{Bind} score³¹. Analysis of MM/GBSA found higher net negative binding free energy values for the five selected and control compounds in complex with MCM7 protein. Thus, the remaining compound has skipped and the five selected compounds CID_387447, CID_5330286, CID_208908, CID_3062316, and CID_987419 along with TSA (CID_444732) has been taken for further evaluation through the DFT based calculation.

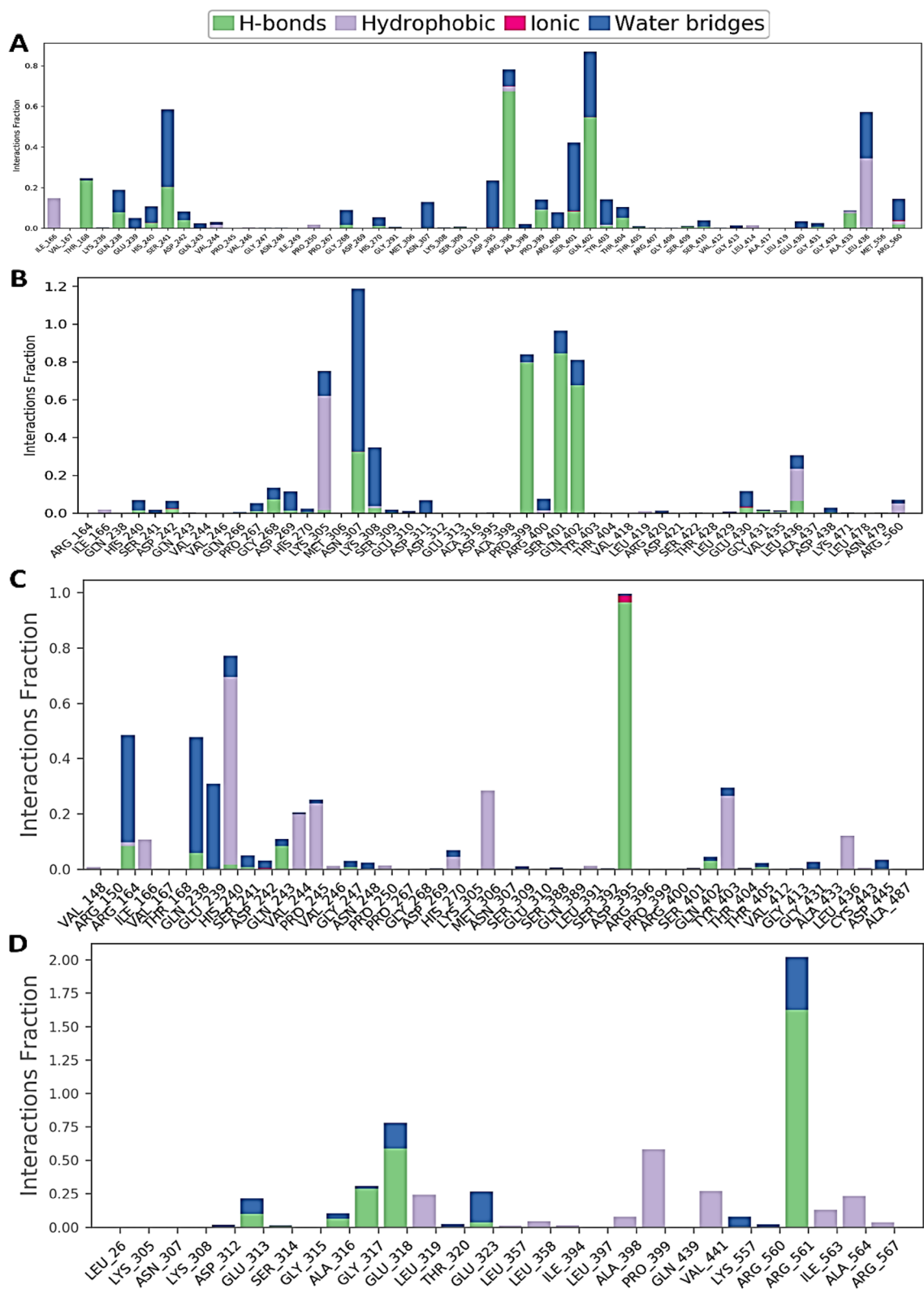


Figure 9. Graphs exhibiting the information about the protein–ligand interaction by 500 ns time of MD simulation. Herein, representing the compounds (A) CID_387447, (B) CID_3062316, (C) CID_9874191, and (D) CID_444372.

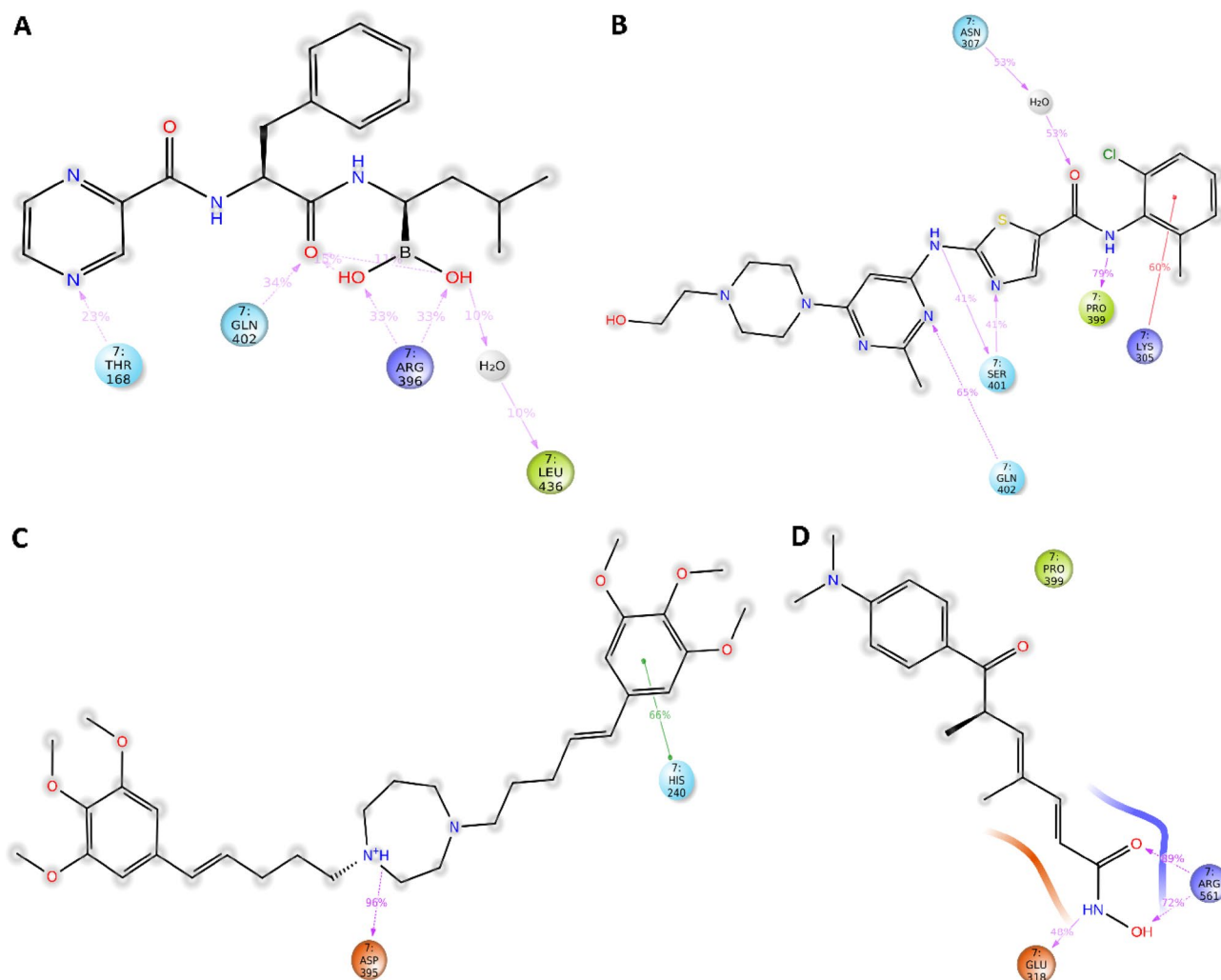


Figure 10. Graph exhibiting the information about the ligand–protein contacts, after 500 ns simulation. In this place, (A) CID_387447, (B) CID_3062316, (C) CID_9874191, and (D) CID_444732 (control).

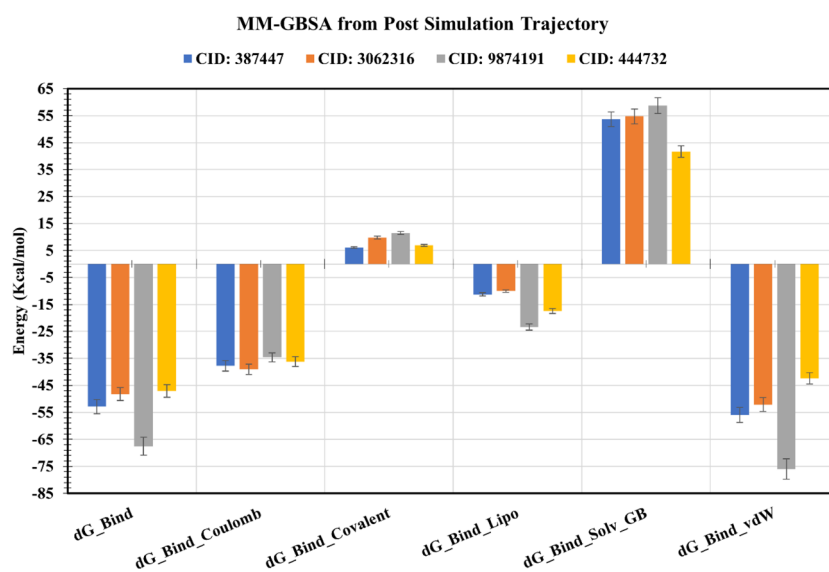


Figure 11. Representing different energy components and net MM/GBSA binding free energy (kcal/mol) and standard deviation values generated from extracted snapshots of MCM7 protein in complex with selected compounds.

Frontier orbitals study or DFT calculation is one of the essential methods determining the pharmacological properties of various small molecules^{31,32}. HOMO and LUMO help to determine and understand the chemical reactivity and kinetic stability of small molecules. The difference between HOMO and LUMO energy is known as HOMO–LUMO gap energy corresponds to the electronic excitation energy. The compound that has the greater orbital gap energy, tends to be energetically more unfavorable to undergo a chemical reaction and can be called bioactive^{33,34}. Moreover, gap energy also correlates with the hardness and softness properties of a molecule³⁵. The DFT calculations were carried out for all the three best ligand molecules compared to the TSA control ligand. Thus, two molecules have been eliminated during the different stages of screening and the remaining three compounds CID_387447, CID_3062316, and CID_9874191 with TSA (CID_444732) were selected for ADMET analysis.

The RO5 demonstrated the drug-like properties of the selected compounds^{36,37}. All the three compounds CID_387447, CID_3062316, and CID_9874191 with control TSA ligand (CID_444732) were found to follow the five Lipinski's rules of drug-likeness properties. The compound with good ADME properties has been further evaluated through the toxicity properties to measure the harmful effect on humans or animals³⁸. Analysis of toxicity found no or less toxicity of the selected three compounds CID_387447, CID_3062316, and CID_9874191 along with TSA.

Molecular dynamics simulation is used to confirm the stability of a protein in a complex with ligands^{30,38}. Also, it can determine the stability and rigidity of protein–ligand complexes in a specific artificial environment like the body³⁰. The RMSD values of the complex systems indicate the best stability of the compounds and RMSF values measure mean fluctuation that determines the compactness of the protein–ligand complex³⁹. The compounds CID_387447 and CID_3062316 showed lowest RMSD and RMSF values except CID_9874191 compared to CID_444732 control compound. The center of mass from the protein C and N terminals tests the stability of the protein structure and gives a broader understanding of protein folding characteristics for Rg calculated⁴⁰. The lower Rg value means that high compactness and the larger value displayed the disassociation of the compounds from the protein and all the compounds except CID_9874191 showed better Rg value. The larger SASA value indicates the less stable structure whereas the lower value means the tightly contracted complex of water molecules and amino acid residues⁴¹. Further, the evaluation of SASA values, hydrogen bond interaction, protein–ligand contact, ligand–protein contact found diverse results, therefore the compound CID_9874191 has been eliminated.

According to the researchers, Dasatinib (CID_3062316) inhibitors could be used to treat lung cancer⁴², gastrointestinal stromal tumors, prostate cancer, multiple myeloma, and sarcomas⁴³. Therefore, Bortezomib (CID_387447) is a reversible proteasome inhibitor that impacts the ubiquitin–proteasome pathway to kill cancer cells, and proteasome inhibition modifies the transcriptional expression of many target genes^{44,45}. The K-7174 (CID_9874191) could be a promising treatment option for chronic disease anemia⁴⁶. The Dasatinib and Bortezomib drugs would be theoretically stable and capable of generating an effective inhibition response to MCM7 protein. Further assessment through various lab-based trial methods can assist with deciding the action of the compound that will give options in contrast to human cancer immunotherapy.

Conclusion

An integrative protein–drug interaction, molecular docking, ADMET, QM calculation, MD simulation, and MM-GBSA approaches revealed CID_387447 and CID_3062316, as potential drug candidates that will help to inhibit the activity of the MCM7 protein against human cancer. Further assessment through various lab-based trial methods can assist with deciding the action of the compound that will give options in contrast to human cancer immunotherapy.

Methods

In this study, we used the Linux (Ubuntu-20.04.1 LTS) operating system and Intel Core i7-10700 K processor CPU, 3200 MHz DDR4 RAM and RTX 3080 DDR6 8704 CUDA core GPU. Computational molecular docking, MM-GBSA, and DFT calculation were generated by Glide, Prime, Jaguar, and Maestro. The molecular dynamic simulation was performed by using the Desmond module of Schrödinger Suite 2020-3.

Protein structure retrieval and preparation. The crystal structure of human MCM7 protein (PDB code: 6XTY) was downloaded from RCSB Protein Data Bank (PDB)⁴⁷. MCM7 protein was co-crystallized with the MCM protein family, therefore the target MCM7 protein was separated and removed the water, metal ions, cofactors, other molecules, and other proteins by the Maestro v-12.5 of Schrödinger Suite 2020-3⁴⁸. The MCM7 protein was initially processed and prepared by the protein preparation wizard (Prep Wizard) of Schrödinger Suite 2020-3⁴⁹. The prepared protein has been further utilized for molecular docking and other experiments.

Compounds identification and preparation. The MCM7 protein and its correspondence compounds interaction were performed by the NetworkAnalyst v3.0 server⁵⁰ and the protein–drugs interaction network was redesigned using the Cytoscape v3.7⁵¹. Therefore, this interacting chemical compound's structure has been searched on the PubChem website⁵² and retrieved in SDF file format. Finally, chemical compounds were processed, refined, and prepared by the LigPrep v-55139⁵³ for molecular docking, MM-GBSA, and DFT calculation.

Binding affinity calculation and MM-GBSA analysis. The desired protein–compounds binding scores have been evaluated by the molecular docking approach⁵⁴. The best binding score of compounds and protein interaction has been analyzed and visualized by using the Glide v-8.8 and Maestro v-12.5.139 respectively^{55,56} (Schrödinger packaged 2020–3) tool⁵⁷. The binding site position of the protein has been determined by reference

ligand active site and a grid box corresponding to the binding site position has been generated. A grid box with a box shape $X = 25.086 \text{ \AA}$, $Y = 25.086 \text{ \AA}$, $Z = 25.086 \text{ \AA}$ has been set for molecular docking simulation. Nowadays most of the lead compounds identified in the CADD process are based on the docking score (protein-chemical compounds binding scores), which does not always provide an accurate and constant score. Therefore, validation of the docking process through different energy calculation methods can provide the reliability of the methods⁵⁸. MM-GBSA (molecular mechanics-generalized born surface area) was calculated to calculate the bind-free energy of ligands and validated the docking process. The MM-GBSA score was predicted by Prime MMGBSA v-3.0⁵⁹. We selected the OPLS_2005 force field⁶⁰ and other parameters were defaults. From this, the docking score and MM-GBSA score were obtained as a control to compare the value with newly screened drugs. Binding interactions, residues, and binding free energy involved in the interacting plane were analyzed with Maestro v-12.5.139⁴⁸.

QM (quantum mechanical) calculation. Conformation analysis of a ligand to the binding site of a protein is an essential part to identify potential active conformation, binding affinity, and strain discipline associated with the binding mechanism. This type of binding possess can be achieved through the calculation of minimum energy conformation and structural optimization, which is dependent on the solution phase and associated gas-phase energy. The classical molecular mechanics (MM) process is unable to describe the process properly due to the presentation of metal ions in a ligand–protein complex system⁶¹. The ligand was minimization by QM (quantum mechanical) calculations using the density functional theory (DFT), these performed by the Jaguar v-10.9⁶². The DFT was treated by B3LYP (Becke exchange functional⁶³, which combined Lee, Yang, and Parrs (LYP)⁶⁴ correlation functional) in conjunction with 6-31G(d,p) basis sets. For this reason, after the highest docking score and MM-GBSA calculated ligand was selected for DFT or QM calculation. In DFT calculation, we analyzed frontier molecular orbitals namely highest occupied molecular orbitals (HOMOs), lowest unoccupied molecular orbitals (LUMOs), and their energy gap difference. When HOMOs energy values show the ability of a ligand molecule to donate electrons, LUMOs energies propose the capability of a ligand molecule to accept electrons from the protein. The frontier energies (ϵ) of HOMOs and LUMOs were used to measure the hardness and softness of selected compounds. The hardness (η) and softness (s) of the drugs were measured performed by the Parr and Pearson interpretation equation⁶⁵ and Koopmans theorem equation⁶⁶. The hardness value determines how the atom resists the charge transfer to another atom or metal surface. The ability of an atom to receive electrons is measured by the softness value. The following Equations can be used to measure the chemical hardness (1) and softness (2).

$$\text{Hardness } (\eta) = (I - A)/2 \quad (1)$$

$$\text{Softness } (S) = 1/\eta \quad (2)$$

In Equation, I refer to the ionization potential ($-E_{\text{HOMO}}$). Also, A denotes the electron affinity ($-E_{\text{LUMO}}$). According to the above-mentioned Equation, the smaller value of hardness means more reactivity and vice versa. Whereas S is the ability of an atom to receive electrons and η is the hardness.

Visualization of selected chemical compounds-MCM7 protein. The constructed protein-compounds interaction network was redesigned using the Cytoscape v3.7⁵¹. The docking binding interactions and ligand DFT calculation figure were generated and visualized with Maestro v-12.5.139⁴⁸. The docking score, MM-GBSA, and MD simulation graphs have been designed in GraphPad Prism 8 and Microsoft Excell365.

Molecular features of the selected chemical compounds. The drug design and development process involves the assessment of ADME/T (absorption, distribution, metabolism, excretion, and toxicity) to identify molecules with the highest pharmacokinetics properties and can be an effective drug¹⁶. The integrity and efficiency of compounds should be described through pharmacokinetic and toxicity properties in the early stages of drug design. To assess the early-stage pharmacokinetic properties of our chosen compounds, the SwissADME (<http://www.swissadme.ch/>) server was used in study⁶⁷. The SwissADME server, a free web-based tool that can analyze small molecule pharmacokinetics and drug-likeness properties. Toxicity evaluation is the crucial stage in the development and design of a drug. Therefore, the toxicity of the selected compounds has been evaluated through the pkCSM (<http://biosig.unimelb.edu.au/pkcsm/>) server⁶⁸.

Molecular dynamics simulation. The complex structure of the selected candidate compounds was evaluated using 500 ns MD simulation to evaluate their binding stability to the desired protein to the active site cavity of the protein⁶⁰. The MD simulation of the receptor-ligand complex whereas performed using the 'Desmond v6.3 Program' in Schrödinger 2020-3 under Linux framework to evaluate the thermodynamic stability of the receptor-ligand complexes³⁰. To solve the system, a predetermined TIP3P water model was used, with an orthorhombic periodic boundary box form with a box distance of 10 Å assigned to both sides to retain a specific volume. Boundary conditions box volume was initially calculated as 910,697 Å³ (CID_9874191); 910,706 Å³ (CID_387447), 910,685 Å³ (CID_3062316) and 910,704 Å³ (control), respectively. Na⁺ and Cl⁻ ions were used to neutralize the system to reach a 0.15 M molar salt concentration. The CID_9874191, CID_387447, CID_3062316, and control ligands was neutralizing the 77, 9, 78, and 78 Na⁺ and 69, 7, 69, and 69 Cl⁻ ions, respectively. After constructing the solvated system containing protein in complex with the ligand, the system has been minimized and relaxed using the default protocol introduced within the Desmond module with OPLS_2005 force field parameters⁶⁰. The Nose–Hoover temperature coupling and isotropic scaling method were

used to keep NPT ensembles at 310 K and one atmospheric (1.01325 bar) pressure, followed by 100 PS recording intervals with an energy of 1.2. The highest temperature of a Newtonian blood fluid flow has been recorded to be 310.0045 K, while the maximum temperature of a non-Newtonian blood fluid flow has been reported to be 310.007 K⁶⁹. As a result, the 310 K temperature has the strongest impact on system⁷⁰. Total number of atoms was calculated as 84,214 (CID_9874191), 86,782 (CID_387447), 84,178 (CID_3062316), and 84,187 (control).

Maestro v-12.5 was used to make all snapshots of MD simulation. The root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), protein-ligand contacts (P-L contact), ligand-protein contacts (L-P contact), and hydrogen bond interaction were used to evaluate the stability of the complex structure based on the 500 ns trajectory performance using the simulation interaction diagram (SID) of Desmond module v6.3.

MM-GBSA analysis from post molecular dynamic simulation trajectory. For calculating the binding free energy of ligands to the macromolecules “molecular mechanics generalized Born surface area” (MM-GBSA) methods have become popular methods⁷¹. The MM-GBSA has been used to estimate the binding free energy of the compounds by using the Maestro package that incorporated in Schrödinger (Release 2020-3) by Using default parameters.

Received: 23 July 2021; Accepted: 14 January 2022

Published online: 27 January 2022

References

- Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA* <https://doi.org/10.3322/caac.21660> (2021).
- Karim, M. A. *et al.* A Multi-omics analysis of bone morphogenetic protein 5 (BMP5) mRNA expression and clinical prognostic outcomes in different cancers using bioinformatics approaches. *Biomedicines* <https://doi.org/10.3390/biomedicines8020019> (2020).
- Omran, A. R. The epidemiologic transition A theory of the epidemiology of population change. *Milbank Meml. Fund Q.* **49**, 509–538 (1971).
- Samad, A., Jafar, T. & Rafi, J. H. Identification of angiotensin-converting enzyme 2 (ACE2) protein as the potential biomarker in SARS-CoV-2 infection-related lung cancer using computational analyses. *Genomics* **112**, 4912–4923. <https://doi.org/10.1016/j.ygeno.2020.09.002> (2020).
- Alam, R. *et al.* A systematic analysis of ATPase Cation transporting 13A2 (ATP13A2) transcriptional expression and prognostic value in human brain cancer. *Biomed. Signal Process. Control* **71**, 103183. <https://doi.org/10.1016/j.bspc.2021.103183> (2022).
- Tye, B. K. MCM proteins in DNA replication. *Annu. Rev. Biochem.* **68**, 649–686. <https://doi.org/10.1146/annurev.biochem.68.1.649> (1999).
- Maiorano, D., Lutzmann, M. & Méchali, M. MCM proteins and DNA replication. *Curr. Opin. Cell Biol.* **18**, 130–136. <https://doi.org/10.1016/j.ceb.2006.02.006> (2006).
- Liang, Z. *et al.* Simvastatin suppresses the DNA replication licensing factor MCM7 and inhibits the growth of tamoxifen-resistant breast cancer cells. *Sci. Rep.* **7**, 41776. <https://doi.org/10.1038/srep41776> (2017).
- Ren, B. *et al.* MCM7 amplification and overexpression are associated with prostate cancer progression. *Oncogene* **25**, 1090–1098. <https://doi.org/10.1038/sj.onc.1209134> (2006).
- Enemark, E. J. & Joshua-Tor, L. On helicases and other motor proteins. *Curr. Opin. Struct. Biol.* **18**, 243–257. <https://doi.org/10.1016/j.sbi.2008.01.007> (2008).
- Ibarra, A., Schwob, E. & Méndez, J. Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. *Proc. Natl. Acad. Sci. USA* **105**, 8956. <https://doi.org/10.1073/pnas.0803978105> (2008).
- Samad, A. *et al.* Computational assessment of MCM2 transcriptional expression and identification of the prognostic biomarker for human breast cancer. *Heliyon* **6**, e05087. <https://doi.org/10.1016/j.heliyon.2020.e05087> (2020).
- Honeycutt, K. A. *et al.* Deregulated minichromosomal maintenance protein MCM7 contributes to oncogene driven tumorigenesis. *Oncogene* **25**, 4027–4032. <https://doi.org/10.1038/sj.onc.1209435> (2006).
- Lau, K. M. *et al.* Minichromosome maintenance proteins 2, 3 and 7 in medulloblastoma: Overexpression and involvement in regulation of cell migration and invasion. *Oncogene* **29**, 5475–5489. <https://doi.org/10.1038/ncr.2010.287> (2010).
- Marnierides, A. *et al.* Immunohistochemical expression and prognostic significance of CCND3, MCM2 and MCM7 in Hodgkin lymphoma. *Anticancer Res.* **31**, 3585–3594 (2011).
- Zhou, W., Wang, Y., Lu, A. & Zhang, G. Systems pharmacology in small molecular drug discovery. *Int. J. Mol. Sci.* **17**, 246. <https://doi.org/10.3390/ijms17020246> (2016).
- Pokhrel, S. *et al.* Spike protein recognizer receptor ACE2 targeted identification of potential natural antiviral drug candidates against SARS-CoV-2. *Int. J. Biol. Macromol.* **191**, 1114–1125. <https://doi.org/10.1016/j.ijbiomac.2021.09.146> (2021).
- Liang, P. H. *et al.* Novel five-membered iminocyclitol derivatives as selective and potent glycosidase inhibitors: New structures for antivirals and osteoarthritis. *ChemBioChem* **7**, 165–173. <https://doi.org/10.1002/cbic.200500321> (2006).
- Wichapong, K., Nueangaudom, A., Pianwanit, S., Sippl, W. & Kokpol, S. Identification of potential hit compounds for Dengue virus NS2B/NS3 protease inhibitors by combining virtual screening and binding free energy calculations. *Trop. Biomed.* **30**, 388–408 (2013).
- Opo, F. A. D. M. *et al.* Structure based pharmacophore modeling, virtual screening, molecular docking and ADMET approaches for identification of natural anti-cancer agents targeting XIAP protein. *Sci. Rep.* **11**, 4049. <https://doi.org/10.1038/s41598-021-83626-x> (2021).
- Liu, Z. *et al.* Identification of hub genes and small-molecule compounds in medulloblastoma by integrated bioinformatic analyses. *PeerJ* **8**, e8670. <https://doi.org/10.7717/peerj.8670> (2020).
- Zhang, X.-F. *et al.* Novel biomolecule lycopene-reduced graphene oxide-silver nanoparticle enhances apoptotic potential of trichostatin A in human ovarian cancer cells (SKOV3). *Int. J. Nanomed.* **12**, 7551–7575. <https://doi.org/10.2147/IJN.S144161> (2017).
- Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* **26**, 795–797. <https://doi.org/10.1038/nbt1418> (2008).
- Hassanpour, S. H. & Dehghani, M. Review of cancer from perspective of molecular. *J. Cancer Res. Pract.* **4**, 127–129. <https://doi.org/10.1016/j.jcpr.2017.07.001> (2017).
- Tabassum, A. *et al.* Transporter associated with antigen processing 1 (TAP1) expression and prognostic analysis in breast, lung, liver, and ovarian cancer. *J. Mol. Med. (Berl)* **99**, 1293–1309. <https://doi.org/10.1007/s00109-021-02088-w> (2021).

26. Meacham, C. E. & Morrison, S. J. Tumour heterogeneity and cancer cell plasticity. *Nature* **501**, 328–337. <https://doi.org/10.1038/nature12624> (2013).
27. Fisher, R., Pusztai, L. & Swanton, C. Cancer heterogeneity: Implications for targeted therapeutics. *Br. J. Cancer* **108**, 479–485. <https://doi.org/10.1038/bjc.2012.581> (2013).
28. Luo, J.-H. Oncogenic activity of MCM7 transforming cluster. *World J. Clin. Oncol.* **2**, 120–124. <https://doi.org/10.5306/wjco.v2.i2.120> (2011).
29. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234. <https://doi.org/10.1007/s10822-013-9644-8> (2013).
30. Bharadwaj, S. *et al.* Exploration of natural compounds with anti-SARS-CoV-2 activity via inhibition of SARS-CoV-2 Mpro. *Brief Bioinform.* **22**, 1361–1377. <https://doi.org/10.1093/bib/bbaa382> (2021).
31. Sarkar, B. *et al.* Identification of the most potent acetylcholinesterase inhibitors from plants for possible treatment of Alzheimer's disease: A computational approach. *Egypt. J. Med. Hum. Genet.* **22**, 10. <https://doi.org/10.1186/s43042-020-00127-8> (2021).
32. Bouback, T. A. *et al.* Pharmacophore-based virtual screening, quantum mechanics calculations, and molecular dynamics simulation approaches identified potential natural antiviral drug candidates against MERS-CoV S1-NTD. *Molecules* **26**, 4961. <https://doi.org/10.3390/molecules26164961> (2021).
33. Hoque, M. M., Halim, M. A., Sarwar, M. G. & Khan, M. W. Palladium-catalyzed cyclization of 2-alkynyl-N-ethanoyl anilines to indoles: Synthesis, structural, spectroscopic, and mechanistic study. *J. Phys. Org. Chem.* **28**, 732–742. <https://doi.org/10.1002/poc.3477> (2015).
34. Zhan, C.-G., Nichols, J. A. & Dixon, D. A. Ionization potential, electron affinity, electronegativity, hardness, and electron excitation energy: Molecular properties from density functional theory orbital energies. *J. Phys. Chem. A* **107**, 4184–4195. <https://doi.org/10.1021/jp0225774> (2003).
35. Ayers, P. W., Parr, R. G. & Pearson, R. G. Elucidating the hard/soft acid/base principle: A perspective based on half-reactions. *J. Chem. Phys.* **124**, 194107. <https://doi.org/10.1063/1.2196882> (2006).
36. Lipinski, C. A. Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov. Today Technol.* **1**, 337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007> (2004).
37. Pollastri, M. P. Overview on the rule of five. *Curr. Protoc. Pharmacol.* **49**, 91211–191218. <https://doi.org/10.1002/0471141755.ph0912s49> (2010).
38. Aljahdali, M. O., Molla, M. H. & Ahammad, F. Compounds identified from marine mangrove plant (*Avicennia alba*) as potential antiviral drug candidates against WDSV, an in-silico approach. *Mar. Drugs* **19**, 253. <https://doi.org/10.3390/md19050253> (2021).
39. Krupanidhi, S. *et al.* Screening of phytochemical compounds of *Tinospora cordifolia* for their inhibitory activity on SARS-CoV-2: An in silico study. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1787226> (2020).
40. Baildya, N., Khan, A. A., Ghosh, N. N., Dutta, T. & Chattopadhyay, A. P. Screening of potential drug from *Azadirachta indica* (Neem) extracts for SARS-CoV-2: An insight from molecular docking and MD-simulation studies. *J. Mol. Struct.* **1227**, 129390. <https://doi.org/10.1016/j.molstruc.2020.129390> (2021).
41. Mahmud, S. *et al.* Computational discovery of plant-based inhibitors against human carbonic anhydrase IX and molecular dynamics simulation. *J. Biomol. Struct. Dyn.* **39**, 2754–2770. <https://doi.org/10.1080/07391102.2020.1753579> (2021).
42. Zhang, M. *et al.* Dasatinib inhibits lung cancer cell growth and patient derived tumor growth in mice by targeting LIMK1. *Front. Cell Dev. Biol.* <https://doi.org/10.3389/fcell.2020.556532> (2020).
43. Gnoni, A., Marech, I., Silvestris, N., Vacca, A. & Lorusso, V. Dasatinib: An anti-tumour agent via Src inhibition. *Curr. Drug Targets* **12**, 563–578. <https://doi.org/10.2174/138945011794751591> (2011).
44. Cacan, E. & Ozmen, Z. C. Regulation of Fas in response to bortezomib and epirubicin in colorectal cancer cells. *J. Chemother.* **32**, 193–201. <https://doi.org/10.1080/1120009x.2020.1740389> (2020).
45. Roccaro, A. M., Vacca, A. & Ribatti, D. Bortezomib in the treatment of cancer. *Recent Pat. Anti-Cancer Drug Discov.* **1**, 397–403. <https://doi.org/10.2174/157489206778776925> (2006).
46. Fujiwara, T. *et al.* A low-molecular-weight compound K7174 represses hepcidin: Possible therapeutic strategy against anemia of chronic disease. *PLoS ONE* **8**, e75568. <https://doi.org/10.1371/journal.pone.0075568> (2013).
47. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242. <https://doi.org/10.1093/nar/28.1.235> (2000).
48. Schrödinger Release 2020-3: Maestro, S. (LLC, 2020).
49. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **27**, 221–234. <https://doi.org/10.1007/s10822-013-9644-8> (2013).
50. Zhou, G. *et al.* NetworkAnalyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* **47**, W234–W241. <https://doi.org/10.1093/nar/gkz240> (2019).
51. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504. <https://doi.org/10.1101/gr.1239303> (2003).
52. Kim, S. *et al.* PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res.* **49**, D1388–D1395. <https://doi.org/10.1093/nar/gkaa971> (2021).
53. Schrödinger Release 2020-3: LigPrep, S. (LLC, 2020).
54. Sándor, M., Kiss, R. & Keseru, G. M. Virtual fragment docking by Glide: A validation study on 190 protein-fragment complexes. *J. Chem. Inf. Model.* **50**, 1165–1172. <https://doi.org/10.1021/ci1000407> (2010).
55. Friesner, R. A. *et al.* Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *J. Med. Chem.* **49**, 6177–6196. <https://doi.org/10.1021/jm051256o> (2006).
56. Halgren, T. A. *et al.* Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750–1759. <https://doi.org/10.1021/jm030644s> (2004).
57. Shivakumar, D., Harder, E., Damm, W., Friesner, R. A. & Sherman, W. Improving the prediction of absolute solvation free energies using the next generation OPLS force field. *J. Chem. Theory Comput.* **8**, 2553–2558. <https://doi.org/10.1021/ct300203w> (2012).
58. Lyne, P. D., Lamb, M. L. & Saeh, J. C. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J. Med. Chem.* **49**, 4805–4808. <https://doi.org/10.1021/jm060522a> (2006).
59. Schrödinger Release 2020-3: Prime, S. (LLC, 2020).
60. Ahammad, F. *et al.* Pharmacoinformatics and molecular dynamics simulation-based phytochemical screening of neem plant (*Azadirachta indica*) against human cancer by targeting MCM7 protein. *Brief. Bioinform.* <https://doi.org/10.1093/bib/bbab098> (2021).
61. Friesner, R. A. & Guallar, V. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu. Rev. Phys. Chem.* **56**, 389–427. <https://doi.org/10.1146/annurev.physchem.55.091602.094410> (2005).
62. Bochevarov, A. D. *et al.* Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quantum Chem.* **113**, 2110–2142. <https://doi.org/10.1002/qua.24481> (2013).
63. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100. <https://doi.org/10.1103/physreva.38.3098> (1988).

64. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789. <https://doi.org/10.1103/physrevb.37.785> (1988).
65. Pearson, R. G. Absolute electronegativity and hardness correlated with molecular orbital theory. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8440–8441. <https://doi.org/10.1073/pnas.83.22.8440> (1986).
66. Parr, R. G. *Horizons of Quantum Chemistry* 5–15 (Springer, 1980).
67. Daina, A., Michielin, O. & Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **7**, 42717. <https://doi.org/10.1038/srep42717> (2017).
68. Pires, D. E., Blundell, T. L. & Ascher, D. B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* **58**, 4066–4072. <https://doi.org/10.1021/acs.jmedchem.5b00104> (2015).
69. Baral, K., Adhikari, P., Jawad, B., Podgornik, R. & Ching, W. Y. Solvent effect on the structure and properties of RGD peptide (1FUV) at body temperature (310 K) using ab initio molecular dynamics. *Polymers* **13**, 3434. <https://doi.org/10.3390/polym13193434> (2021).
70. Mahdavi, M., Fattahi, A., Tajkhorshid, E. & Nouranian, S. Molecular insights into the loading and dynamics of doxorubicin on pegylated graphene oxide nanocarriers. *ACS Appl. Biol. Mater.* **3**, 1354–1363. <https://doi.org/10.1021/acsabm.9b00956> (2020).
71. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **10**, 449–461. <https://doi.org/10.1517/17460441.2015.1032936> (2015).

Acknowledgements

We extend our gratitude towards the Biological Solution Centre (<https://biosolcentre.org/>) for providing software and hardware support.

Author contributions

A.S. designed the project, generating the data, analyzing the data, and wrote the manuscript; M.S.R. designed the project, supervision, and reviewed the draft manuscript; M.A.H. addressed all the concerned raised by the reviewer and designed the study. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-05621-0>.

Correspondence and requests for materials should be addressed to M.A.H. or M.S.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022